Spatial Perception in Virtual Environments: Evaluating an Architectural Application

Daniel Henry, Research Assistant and Dr. Tom Furness, Director
Human Interface Technology Laboratory, FU-20
University of Washington
Seattle, WA 98195

ABSTRACT

An experiment was conducted to compare and explore the relationship between the way people perceive real and virtual spaces. Twenty-four architects toured either a real museum gallery or a real-time computer generated model of the same gallery under one of three increasingly inclusive viewing conditions i.e. looking at a monitor, viewing through stereoscopic head-mounted displays without and with head-position tracking. Subjects were asked to perform spatial dimension, orientation and evaluation tasks. The most significant results indicated that subjects consistently underestimated the dimensions of the gallery in all three computer simulation conditions when compared to touring the real gallery. Furthermore, the most inclusive viewing condition yielded underestimates for spatial dimensions which were significantly greater than the other two simulation conditions.

KEYWORDS

Virtual Environment, Architectural Simulation, Spatial Perception, Orientation, Field of View.

INTRODUCTION

Of the many potential applications for virtual environment technologies, few are perhaps as obvious as the simulation of architectural spaces. The real-time spatial representation of graphical data which defines virtual environments makes them perfect tools for simulating our physical environment which itself is also spatial and perceived in real-time. The simulation of architectural spaces using virtual environments also satisfies a real need in the architectural community, that of improving communication between the designer who relies on architectural drawings, and the client whose best understanding of the design occurs when actually physically present within the finalized space. Virtual environments would allow architects and their clients to make cost saving decisions early in the design process and would increase the likelihood for clients to be pleased with the finished design.

Presently, clients and architects can discuss design issues by peering through or down into scale models. The appeal of this medium is that it is spatial, as is the finished product. The disadvantage with scale models is that clients have to "project" themselves into the model. They have to imagine themselves looking and moving about the model. This operation is prone to errors in spatial perceptions because of the required change in scale.

To reduce potential misperceptions, moveable miniaturized cameras have been introduced into the model. The images captured by the camera are then displayed on a television monitor. This solution places the clients' viewpoint at the appropriate viewing locations in the model. Similar to the miniature camera tour are real-time computer generated models which are also displayed on monitors. Walkthrough™ is one such program, allowing clients and architects to explore computer generated model of a project interactively. Both solutions facilitate the appreciation of the space by positioning the viewpoint of the viewer at the appropriate eye height, thereby diminishing an important source of error. However, the display of the spaces on relatively small monitors still requires users to interpret scale. Furthermore, viewers now have to make judgments about 3-D space while viewing it in 2 1/2-D.

Problems of accurate eye height position, scale adjustment and the need for 3-D representation can all be addressed with virtual environment technologies. Head-position tracking automatically updates the images so as to place the user's viewpoint correctly within the scene. Due to the proximity of the eyes to the display in the HMDs (head-mounted display), adjustment for scale is no longer required. And, the illusion of seeing a 3-D environment is conveyed by the stereoscopic displays of images which completely surround the user.

Not surprisingly, numerous research institutions have already begun to explore and expand on the use of virtual environment technology for architectural applications. Refining the resolution of HMDs, increasing the

0-7803-1363-1/93 $3.00 © 1993 IEEE

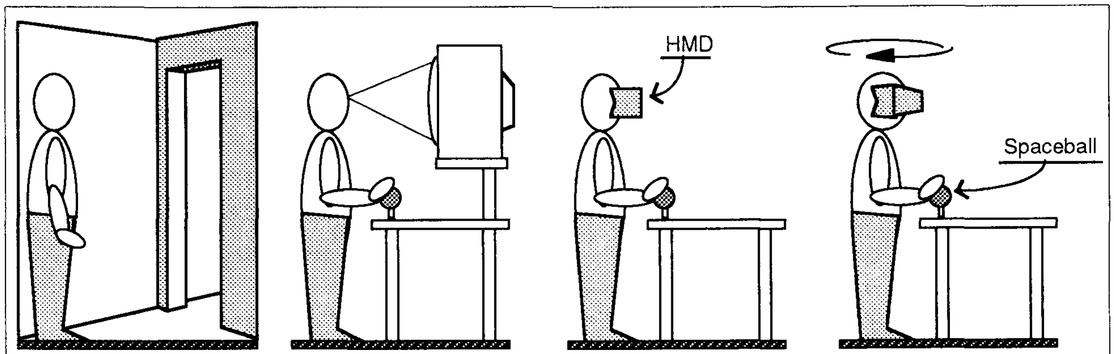
update rate, improving the accuracy and speed of position sensors, improving the quality of the graphics, designing tools for interacting with the model, and developing interaction devices for simulating the act of walking are examples of technology developments meant to improve the use of this tool for architectural applications [1].

Although these technological improvements are well intended, they are all based on the untested assumption that the *basic characteristics of spaces*, that is to say, their shapes, their relative location, and their "feel" are perceived the same way in both virtual and real environments. Everything about the technology would tend to indicate that this is the case. However, this cannot be taken for granted. Architect's clients would probably be unhappily surprised at the outcome of a project if such a tool were to unknowingly lead them to make erroneous perceptions of space. It is therefore imperative to address the following basic questions: how well does this new tool represent architectural spaces? How accurate can the perception of virtual spaces be in predicting the perception of real spaces? What are the weaknesses and strengths of this simulation tool?

The purpose of this study was to explore and determine the accuracy of virtual interfaces in simulating the *basic characteristics of architectural spaces* and to evaluate to what extent our perceptions of virtual and real spaces coincide. By including increasingly inclusive forms of the technology in the study, from monitor viewing described above, to the fully interactive and inclusive interface, it would also be possible to establish what improvement in spatial perception, if any, complete virtual environments would offer over existing architectural representations.

THE EXPERIMENT

The experiment consisted of comparing subjects' spatial perception of an existing architectural space under four viewing conditions (fig.1): (1) touring the real place (*Real condition*), (2) interacting with a computer generated model viewed on a television monitor (*Monitor condition*), (3) viewing the model through stereoscopic HMDs without head-position tracking (*Fixed condition*) and (4) with head-position tracking (*Tracked condition*). The Monitor condition represents the state of technology available to design professionals today. The Tracked condition is considered the "fully immersive" virtual environment. Some of the factors which make the Tracked condition different from the Monitor condition include: total isolation of the peripheral visual field of view (like swimming goggles, the sides of the HMD are opaque), rotation of viewpoint through control of head movement (direct coupling) rather than the less direct use of the joystick, parallax (improved depth perception) and to a lesser degree, lower screen resolution. The Fixed condition was meant to help isolate the multiple factors which distinguish the Monitor and the Tracked conditions.



REAL CONDITION	MONITOR CONDITION	FIXED CONDITION	TRACKED CONDITION
Rotation of viewpoint and Movement controlled with movement of body.	Rotation of viewpoint and Movement controlled with Spaceball.	Rotation of viewpoint and Movement controlled with Spaceball.	Rotation of viewpoint controlled by turning head. Movement controlled with Spaceball.

fig.1 - The four touring conditions

The Equipment - The computer model of the test space was constructed using Alias modeling software. The degree of detail was determined by what could be seen with the resolution of the HMDs. Scale figures were included in the model to simulate the experimenter's presence in the Real condition tour. The generation of graphics in real-time (8 to 10 frames/second) were computed on the Silicon Graphics VGX 320. The images were displayed through the VPL Eyephones, which have a resolution considerably lower than NTSC. The horizontal field of view in the VPL Eyephones is 75° per eye and 90° for both eyes combined (there is a 60° overlap). In the Monitor condition, one of the HMD channels was displayed on a 19" color monitor. The distance between the participant and the monitor was such that the horizontal field of view remained the same across all simulation conditions (90°).

Participants moved their viewpoint through the model with the aid of the SpaceBall (a pressure-sensitive joystick) located in front of them. Movement was constrained in velocity (1.5 m/s for walking) and in direction (no vertical movement). Rotation of the viewpoint in the *yaw* (turning head) and *pitch* (nodding head) directions was controlled by the Spaceball in both the Monitor and Fixed conditions. In the Tracked condition, a Polhemus head-position tracker placed on the HMD measured the rotation of the user's head. This input was gathered by VPL's DataBox. A simple program designed at the H.I.T.Lab interpreted the Spaceball and DataBox data in order to update the images the user was seeing.

The Comparison Site - The study was conducted in the art galleries (without the artwork) of the Henry Art Museum at the University of Washington. A museum space was selected partly because similar spatial studies used museum spaces [2], and partly because the gallery spaces offer a setting which is relatively simple and easy to control: there are no windows, surfaces are orthogonal, lighting is indirect, the flux of people not in the study can be limited and there are few unwanted sounds.

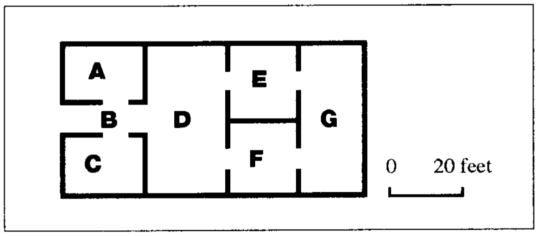


fig. 2 - Plan of museum and labeled spaces.

The ability to control these spatial elements was important because the update rate for real-time interaction in virtual environments is directly affected by the number of polygons (detail) in the model. More complex environments such as exterior scenes and multiple light sources would have resulted in a slower rate of update for the images. Furthermore, additional environmental factors such as crowding, draft and unpleasant sounds are known to affect people's perception of spaces, thereby complicating the interpretation of the results.

The test space consisted of six galleries and one hallway (fig.2). Galleries varied in volume from 17 by 23 feet with 14 foot ceilings to 28 by 47 feet with 20 foot ceilings. All of the galleries were lit with diffused daylight from skylights above, an effect which could be easily reproduced in the computer model. Spaces A, E and G contained one piece of furniture (television or chair) which would be referred to in the tasks and which also served as scale figures in both the real gallery and the computer model.

Subjects - Twenty-four professional architects, architectural graduate students and professors of architecture participated in this study. This homogeneous group was selected because their professional background would guarantee a certain degree of consistency in their spatial perception. They also represent a very large potential user group for this technology, and their inclusion in the experiment is important for the design process.

Design - Because of the lingering effects of spatial memory, the subjects could not be presented a completely random design. Instead, a random block design was selected. Each of the 24 subjects was randomly selected to visit the galleries in <u>one of four</u> viewing conditions; the Monitor, Fixed, Tracked or Real condition. It was hoped that the difficulty of getting significant results with only 6 subjects per block in a random block design would be partially offset by the homogeneity of the subject group.

Tasks - The *basic characteristics of space* were defined as (1) volume shape and dimension, (2) relative location of volumes and (3) the general "feel" of individual volumes. Shape and dimension are dominant characteristics of space [3]. They are largely responsible for the way the space is perceived. To compare perception of dimensions, subjects were asked to estimate the length, width and height of three gallery spaces. It was assumed architect's perception of dimensions would be very accurate because they perform similar tasks in their everyday work.

A second basic characteristic of space is that each has a specific location within a larger environment. The feel of a "sub-space" such as an bedroom in a house, can vary depending on its relative location to the larger space. Our ability to sense this is based on our ability to build an accurate space plan or "mental model" of our position within the larger environment. If our mental model is the same across viewing conditions, then it can be assumed that the relative position of the sub-spaces will also be accurately perceived. The task to test for this required subjects to point at objects they had seen during the tour but which were now out of sight. This is a common "orientation" task [4&5].

In an attempt to confirm the results of the pointing task, participants were also asked to sketch a plan of the spaces after the tour. This is a common method for measuring people's mental model of spaces [6]. The principal reasons for not relying entirely on sketch tasks is that they are difficult to evaluate quantitatively.

While the "feel" of spaces is not a specific characteristic of space, it reflects in a global sense the overall effects of all of the spatial characteristics which make up the space. In this respect, it is perhaps the most important unit of measure for spatial perception. It includes such difficult things to measure as the effect of lighting, color, temperature, as well as other less tangible aspects of space. It was assumed that the sense of "presence" one acquires in a virtual environments would appreciably enhance people's ability to judge these less tangible aspects of spaces. Environmental psychologists have developed a wide range of techniques for measuring people's response to spaces [7]. Techniques such as adjective, mood and noun checklists were used very effectively in similar architectural representation studies [8]. For this study, subjects evaluated the feel of gallery D by using a 13 bi-polar adjective checklist which was drawn from a more complete list developed by Kasmar [9].

Procedure - Participants selected for the Real condition were greeted at a nearby building to fill out a personal profile questionnaire. Then, they were taken to space "B" in the museum. In the case of the computer model touring conditions (Monitor, Fixed and Tracked), more preparation was necessary. It was feared that differences in subjects experience using computer visualization tools could distort results. It was also important to diminish the "novelty factor" of the technology. To alleviate these concerns, each participant was introduced to a 15 minute visit of a pre-test virtual environment, touring it in the same viewing condition they would be using in their tour of the actual test model. After this preparation, the gallery model was displayed and participants also found themselves in space "B".

Next followed a 15 minute tour of the gallery spaces. The experimenter proceeded in guiding the participants to space A where they were asked to estimate the dimensions (width, length and height) of the volume. They were also asked to pay particular attention to the location of the chair in the space because they would be asked to guess its location at a later time. Then they were taken back through B, D, F and into G to estimate the dimensions of that space. They were also asked to point in the direction of the chair they had seen in space A. In the simulation conditions, rather than pointing, participants centered their views in the direction of the chair. Participants were then taken to the other end of G, through E and into D where they performed the dimension task once more. They were then asked to point to the chair in space F. They were also instructed that they would be asked to qualify this space in a questionnaire after the tour. Participants were then taken through B and into C and asked to point to the chair in space G. That concluded the tour. Finally, subjects were invited to fill out a 30 minute questionnaire which included the sketching task and the bi-polar adjective checklist, as well as other questions pertaining to the sense of "presence".

RESULTS

Dimension task - The data for the dimensions of the spaces were divided into two groups: horizontal and vertical dimensions. This was found to be necessary because it was unequally easy for participants to establish the three spatial dimensions. Many participants advanced the fact that interior spaces usually come in standard heights, thereby considerably improving subject's estimates in that dimension (24% of all vertical estimates were "on the nose" compared with 1% for the horizontal dimensions). The standing posture of the subjects also improved their estimates because they could use their vertical height as a scale.

A multi-comparison one factor ANOVA for <u>horizontal</u> distance estimates shows a significant main effect due to display conditions with $p<.0001$ at the 95% confidence interval. The underestimates for dimensions in all of the simulation conditions are significantly different from the Real condition. Results for the Real display condition were quite accurate. Among the simulation conditions, the estimates in the Tracked condition were

significantly smaller than the other two display conditions, according to the Fisher PLSD test. Furthermore, the more conservative Scheffe F-test suggests that the Tracked condition was also significantly different from the Real condition (fig.3).

A multi-comparison one-factor ANOVA for vertical estimates suggests there is a main effect between display types, with p<.0003 at 95% confidence interval. The means for the simulation conditions are all smaller than the mean for the Real condition, as was the case for the horizontal dimension (fig. 3). Both the Fisher PLSD and the Scheffe F-test indicate that the Tracked condition results are significantly smaller than the Monitor and the Real condition. Both tests for individual differences also show that, in the two HMD conditions, Fixed and Tracked, the estimates were significantly smaller than in the Real condition.

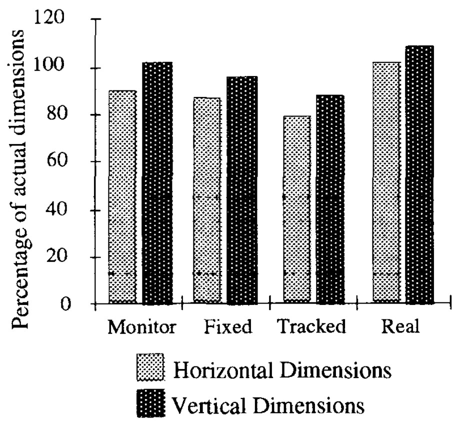
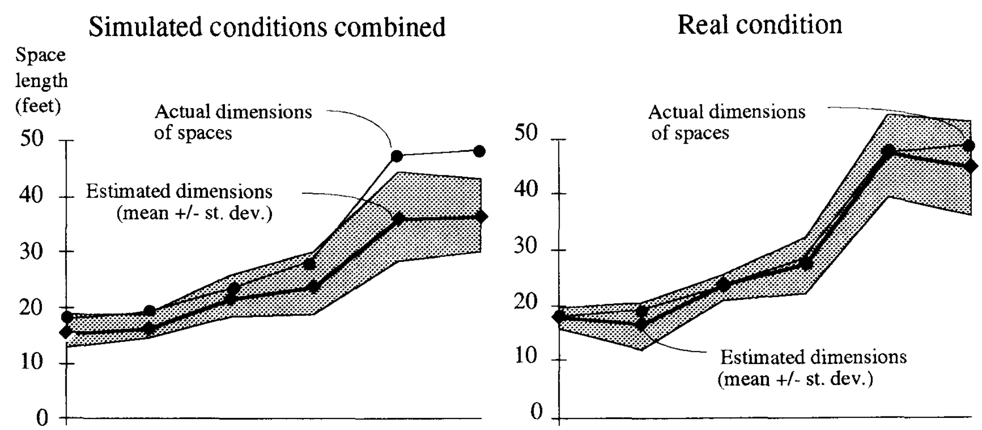


fig. 3. - Average dimension estimates

While the effect of underestimating was more pronounced in the Tracked condition, when combined, estimates in all three simulation conditions were significantly short of the real dimensions (fig. 4). This tendency increased as the spatial dimensions increased. Data for the height estimates did not reveal such a large discrepancy between the simulated and the Real conditions. This might be due to the advantages subjects expressed estimating vertical dimensions.



There are 6 horizontal dimension estimates in the test. Each point represents the averaged estimate of one of the 6 dimensions. They are presented in order of increasing size.

fig.4. Comparison of horizontal dimension estimates with actual dimensions.

Dimension task discussion - One intriguing aspect of these results is that, of the three simulation conditions, the fully immersive Tracked condition yielded the most underestimated spatial dimensions. In the horizontal dimension, these were as much as 20% lower than estimates in the Real condition. This difference cannot be attributed to the HMDs themselves because results in the Fixed condition (which also included viewing through the HMDs) did not show as consistently large underestimates. Rather, it would appear that the source of difference between the Tracked condition and the other two "non-tracked" conditions lies in differences in the process of doing the task. In the non-tracked conditions (Monitor and Fixed), to turn their viewpoint, participants rotated the model into their view. In the Tracked condition, subjects turned their heads (and bodies) in search of the room edges, much as they would in a real space. In the anticipation of upcoming spatial

information, subjects probably turned their eyes towards the edges of the HMDs. As a result, they would have been looking through the edges of the HMDs optics, precisely where the distortion of the optics is greatest.

The findings also suggested that all of the horizontal dimension estimates in the simulation conditions combined were significantly smaller than those in the Real condition. There are several possible explanations. The simplest explanation is a scale problem in the computer model. However, this is not likely because the scale of model in the present system was calculated very accurately. Another explanation is related to the difference in the way people moved through the spaces. As discussed earlier, the act of physically walking is a very important cue for getting a sense of rate of movement and distance [10]. Participants in the real condition physically paced across the spaces. Participants in the simulation conditions had no kinesthetic feedback for their movement. There is little reason to believe however that simulated movement should necessarily make distances consistently appear to be shorter, as the result have shown.

A more probable explanation might be difference in visual field of view between the Real and the simulation conditions. The horizontal visual field of view is about 180° in the Real condition, as compared to 90° in all three simulation conditions. Viewers in the simulation conditions were all deprived of spatial information in their peripheral field of view. The underestimated results would concord with other studies which have also shown that the perception of size and distances diminish as the field of view narrows [11 &12].

In addition, a study by Hagen would explain why underestimates in the Monitor condition were less pronounced than in the other two conditions. The study suggests that estimates of size and distance are more accurate in conditions where the peripheral vision has *at least some* visual stimulus, even if the stimulus has no direct relationship in scale with the information in the fovea [13]. The participants in the Monitor condition could see the rest of the laboratory in their peripheral field of view. It served as a context, or a frame of reference with which to scale images on the screen. Whereas participants viewing through the HMDs had no visual information in the periphery other than the black rubber sidings of the device.

There are of course numerous other differences between the real museum and the simulated one. People walking through the real museum could hear their footsteps. They could feel and smell the qualities of the air, and they probably perceived many other aspects to the space which did not exist in the simulation condition, and yet which played a role in influencing their perceptions [14].

The orientation task - A bias in the different pointing tasks required the normalization of the third task with respect to the first two. This is a common phenomenon in orientation tasks [4]. After normalization, a multi-comparison one factor ANOVA for the angles showed no main effect for display condition, $p<0.17$ at the 95% confidence interval. What is clear from the diagram however is the rather large precision in angle estimates in the Real condition versus the broader distribution of estimates in the simulation conditions (fig.5).

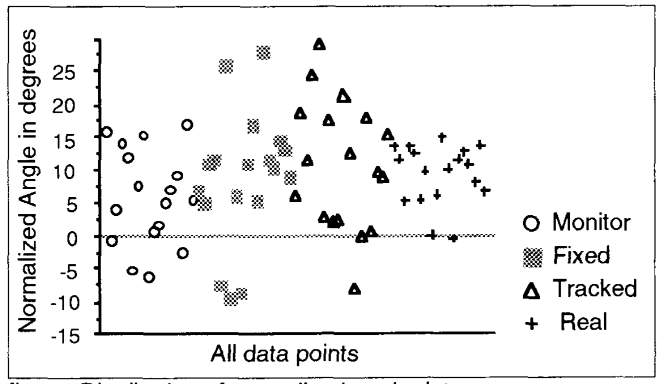


fig. 5 - Distribution of normalized angle data.

Orientation task discussion - There are several reasons which might explain the relative similarity in results across viewing conditions. The first might be the simplicity of the test space as well as the repeated viewing of some of the spaces. Studies have shown that "repeatedly viewing a spatial event increases the accuracy of the observer's spatial representation" [15].

Another important factor was the process of the task. One of the other effects of having a limited field of view in the simulation conditions is that participants could not see as much of the space at once as those in the Real condition. So while participants in the simulation conditions might have known the location of the target, their view of the space was too limited for them to express the direction accurately. As a result, it is unclear whether the limited field of view hampered people's ability to build a cognitive map, or whether it just made the task of identifying their selected direction difficult.

While results are inconclusive, the limited field of view remains a concern in its effect on the development of a mental model of the space. With respect to a restricted field of view, Alfano states that "the overlap of peripheral and fovea information is necessary for veridical perception to occur, ... and that restricting the field of view will interfere with both perception and visuomotor performance" [12].

The "feel" of the spaces - A simple regression was computed with the data from the adjective checklist to measure which of the three simulation condition evoked impressions of the space which were most like those felt in the Real condition. The regression values were as follows: Real and Monitor .88, Real and Fixed .79, and Real and Tracked .91.

The "feel" discussion - The results from the adjective checklist are all relatively similar. A more complex test space as well as a more complete checklist might have generated larger distinctions between the conditions. Nonetheless, the regression values indicate a noticeably greater concordance between the Real condition and the Tracked condition than with the other two simulation conditions. This would seem to indicate that head-position tracking allows people to better judge the feel of simulated spaces. This might support the notion that the greater sense of "presence" felt when viewing virtual environments do make it is much easier to judge the feel of a space. That is afterall the conceptual difference between the Monitor and the Tracked conditions; in the first, one *looks at* a space while in the second, one is *present within* the space.

CONCLUSION

The extent to which distances are perceived to be smaller in the simulated environments is perhaps the most important finding. It would appear that this misperception is due to the well documented size-constancy phenomenon, whereby sizes and distances appear to be smaller when seen through a truncated field of view. The underestimation of distances is large enough to raise concerns about the uses of the tool for doing spatial evaluations. If used with a display configuration similar to the one in this study (HMDs with 90° field of view), they could lead users to make considerable errors in judgments of volume size.

With respect to orientation, the results of this study are unsatisfactory. This is primarily because the design of the task corrupted the results. The limited field of view in the simulation conditions made it difficult for subjects to see which way they were facing. They often found themselves facing a wall which filled their entire field of view. The concern for orientation remains however. Informal observations at the H.I.T.Lab indicate that the limited field of view does hamper people's ability to orient themselves. And while the same might be said of monitor viewing, the more engaging aspect of virtual environments make the experience of being lost more unpleasant. This is certainly a place for further study.

Finally, results from the task of qualifying the space start to suggest that the combination of stereoscopic viewing and head-position tracking together do improve people's perception of spaces. The ability to "feel" as if they are there certainly must enhance their identification with the spaces.

Generally, the most conclusive aspects of this study indicate that the *basic characteristics of spaces* in virtual and real environments are perceived differently with respect to the size of volumes. This discrepancy also appears to be related not to virtual environments per say, but to the truncated field of view which some of the display solutions offer. However, the *inclusive* and *interactive* nature of this tool conveys quite adequately the less tangible notions about how a place feels. These results should steer architectural users of virtual environments towards using forms of displays which offer a wider field of view (such as retro-projection) until further research can be conducted with respect to the perception of distances in HMDs.

FUTURE RESEARCH

It would be important explore the effect of varying fields of view on people's perception of distances and people's sense of orientation. Such a study might deliver a kind of "scale of underestimation" as a function of available field of view. It would also be interesting to study how the physical act of walking affects spatial perception in virtual environments. Interfaces which were not available for this study but which introduce kinesthetic feedback include optical sensors (permitting long range walking) and the treadmill, both developed at U.N.C.

ACKNOWLEDGMENTS

The authors would like to thank the Henry Art Museum Director Richard Andrews as well as his entire staff for their input and the use of their gallery spaces. We also thank professors Woodrow Barfield, Judith Ramey and Robert Sasnoff for their precious support, as well as the entire H.I.T.L. staff. Finally, we extend our thanks to the members of the architectural community who took part in this study with such complete sincerity.

REFERENCES

[1] Brooks, Frederick P. Walkthrough - A dynamic graphics system for simulating virtual buildings. *Interactive 3D Graphics*, Department of Computer Science, University of North Carolina at Chapel Hill, October 23-24, 1986.

[2] Sasanoff, Robert and Bonsteel, David L. *An Investigation of a Televised Image in Simulation of Architectural Spaces*, Architecture Development Series n°. 6, College of Architecture, University of Washington, 1967.

[3] Thiel, Philip. *Towards an Experimential Envirotecture*, (unpublished), College of Architecture and Urban Planning, University of Washington, 1964.

[4] Okabe, Atsuyuki, Oaki, Ken and Hamamoto, Wataru. Distance and direction judgement in a large-scale natural environment: The effects of a winding trail. *Environment and Behavior*, 1986, Vol. 18, n°. 6, 755-772.

[5] Lindberg, Erik and Gärling, Tommy. Aquisition of locational information about reference points during blindfolded and sighted locomotion: Effects of a concurrent task and locomotion paths. *Scandinavian Journal of Psychology*, 1981, Vol. 22, 101-108.

[6] Lynch, Kevin. *The image of the City*, M.I.T press, Cambridge, Massachusetts, 1960.

[7] Craik, Kenneth H. The Comprehension of the Everyday Physical Environment, *Environmental Psychology* (Proshansky, Ed.), Holt, Rinehart and Winston , Inc., 1970.

[8] Bosselmann, Peter and Craik, Kenneth H. "Perceptual Simulations of Environment", *Methods in Environmental and Behavioral Research* (Bechtel, Robert B., Marans, Robert W. and Michelson, William. Ed.), Van Nostrand, New York, 1987.

[9] Kasmar, Joyce Vielhauer. The development of a usable lexicon of environmental descriptors. *Environment and Behavior*, 1970, September, 153-169.

[10] Gibson, James J. *The Ecological Approach to Visual Perception*, Houghton Mifflin Company, Boston, 1979.

[11] Dolezal, Hubert. Living in a World Transformed - *Perceptual and Performatory Adaptation to Visual Distortion*, Academic Press Inc., New York, 1982.

[12] Alfano, Patricia L. and Michel, George F. Restricting the field of view: Perceptual and performance effects. *Perceptual and Motor Skills*, 1990, 70, 35-40.

[13] Hagen, Margaret A., Jones, Rebbeca K. and Reed, Edward S. On a neglected variable in theories of pictorial perception: Truncation of the visual field, *Perception & Psychophysics*, 1978, vol. 23 (4), 326-330.

[14] Canter, David. *The Psychology of Place*, St. Martins Press, New York, 1977.

[15] Allen, Gary L., Siegel, Alexander W., and Rosinski, Richard R. The role of perceptual context in structuring spatial knowledge. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, Vol.4, N°. 6, 617-630.

[16] Passini, Romedi. *Wayfinding in Architecture*, Van Nostrand Reinhold Company Inc., New York, 1984.